Available online at www.sciencedirect.com
SCIENCE DIRECT
Economic Modelling 22 (2005) 391–422
Economic Modelling
www.elsevier.com/locate/econbase

Evidence on structural changes in U.S. time series

Jamel Jouini*, Mohamed Boutahar[1]

GREQAM, Université de la Méditerranée, 2 rue de la Charité, 13002 Marseille, France

Accepted 18 June 2004

Abstract

Some developments in the analysis of structural change models are presented to explore the empirical evidence of the instability by uncovering structural breaks in some U.S. time series. Indeed, the unstable series may be subjected to a meticulous examination allowing to determine changes present in their structure. To that effect, we pursue a methodology composed of different steps and we propose a modelling strategy so as to detect the break dates that can exist in the series. Once the breaks are selected, we attempt to find economic explanations showing why in the chosen dates there are changes in the series. The results indicate that the time series relations have been altered by various important facts and international economic events such as the two Oil-Price Shocks and changes in the International Monetary System. They also show that the presence of high correlation in the data greatly affects the estimation precision of some procedures.

JEL classification: C20

Keywords: Structural change; Break dates; Hypothesis testing; Model selection

1. Introduction

The econometrics and statistics literature holds an important volume of works related to the problem of structural change. For linear regression models, there exist the works of

* Corresponding author. Tel.: +33 4 91 14 07 23; fax: +33 4 91 90 02 27.
E-mail addresses: jouini@ehess.vcharite.univ-mrs.fr (J. Jouini), boutahar@lumimath.univ-mrs.fr (M. Boutahar).
[1] Tel.: +33 4 91 82 90 89; fax: +33 4 91 82 93 56.

0264-9993/$ - see front matter
doi:10.1016/j.econmod.2004.06.003

Quandt (1958) and Chow (1960) who consider tests for structural change for a known single break date. The researches headed for the modelling where this break date is treated as an unknown variable. Indeed, Quandt (1960) extends the Chow test and proposes taking the largest Chow statistic over all possible break dates. In the same context, the most important contributions are those of Andrews (1993) and Andrews and Ploberger (1994) who consider a comprehensive analysis of the problem of testing for structural change. Sen and Srivastava (1975a,b), Hawkins (1977), Worsley (1979, 1986), Srivastava and Worsley (1986) and James et al. (1987) consider tests for mean shifts of normal sequence of variables. The multiple structural changes case receives an increasing attention. Indeed, Yao (1988), Yin (1988) and Yao and Au (1989) study the estimation of the number of mean shifts of variables sequence using the Bayesian information criterion. Liu et al. (1997) consider multiple changes in a linear model estimated by least squares and estimate the number of changes using a modified Schwarz' criterion. Recently, Bai and Perron (1998) consider the estimation of multiple structural shifts in a linear model estimated by least squares. They propose some tests for structural change for the case with no trending regressors and a selection procedure based on a sequence of tests to estimate consistently the number of break points. Bai and Perron (2004) assess via simulations the adequacy of these methods. Indeed, they study the size and power of tests for structural change, the coverage rates of the confidence intervals for the break dates and the relative merits and drawbacks of model selection procedures.

Some international events such as institutional changes, Oil-Price Shocks, change in the Federal Reserve's operating procedures in 1979, etc., may be the source of structural breaks in the series. Perron (1989) carries out standard tests of the unit-root hypothesis against trend-stationary alternatives with a break in the trend occurring at the Great Crash of 1929 or at the Oil-Price Shock of 1973 using the Nelson–Plosser macroeconomic data series as well as a postwar quarterly real gross national product series. His tests reject the null hypothesis of unit root for most of the series if the true data-generating process is that of stationary fluctuations around a trend function that contains one structural break. In the same context, Zivot and Andrews (1992) consider a variation of Perron's tests in which the break date is estimated rather than fixed.

The literature addressing the issue of structural change in stationary time series is relatively important, for example, a process that changes from one stationary process to another. On the other hand, the works focussing on the changes in nonstationary time series are relatively sparse. Recently, there have been works in this context as those of Hansen (1992a). Mankiw and Miron (1986) and Mankiw et al. (1987) find an interesting case of structural change. Indeed, the short-term interest rate has changed from a stationary process to a near random walk since the Federal Reserve System (FRS) was founded at the end of 1914. However, the asymptotic theory on this kind of structural change started to be explored by Chong (2001) who develops a comprehensive asymptotic theory for an autoregressive process of order one with a single structural change of unknown date. More precisely, he examines the case where an AR(1) process changes from a stationary one to a nonstationary one (or the other way around). In each case, he establishes the consistency of the estimators and derives their limiting distributions. He buckles the study by performing some simulation experiments to see how well his asymptotic results match the small-sample properties of the estimators.

In this paper, we attempt to explore the empirical evidence of the instability based on some recent developments in the analysis of structural change models and by pursuing a methodology composed of four steps. The first consists in testing for structural change; in the second step, we estimate the number of breaks; in the third step, we estimate and construct confidence intervals for the break dates and the last one consists in estimating the regression parameters of each regime. We present the various techniques used in each step, and we propose a modelling strategy to illustrate the applicability of these techniques using U.S. time series. We also attempt to find economic explanations showing why in the chosen dates there are changes in the series.

The remainder of the paper is organized as follows. Section 2 presents the structural change model and the estimation method. Section 3 addresses the issue of testing for structural changes. Section 4 presents some model selection procedures. Section 5 considers the problem of estimating the break dates and forming confidence intervals. A few empirical applications and comments are reported in Section 6. The results show that some procedures perform reasonably well since they lead to an adequate number of breaks with locations coinciding with important international economic events such as the two Oil-Price Shocks and the change in the Federal Reserve's operating procedures in 1979. Another feature of substantial importance is that the high correlation in the series affects the results of some procedures. Section 7 contains some concluding comments.

2. Structural change model and estimators

Consider the following structural change model with m+1 regimes:

$$\begin{array}{ll} y_t = \boldsymbol{x}_t'\beta + z_t'\delta_1 + u_t, & t = 1,2,\dots,T_1 \\ y_t = \boldsymbol{x}_t'\beta + z_t'\delta_2 + u_t, & t = T_1+1,\dots,T_2, \\ \vdots & \vdots \\ y_t = \boldsymbol{x}_t'\beta + z_t'\delta_{m+1} + u_t, & t = T_m+1,\dots,T, \end{array} \tag{1}$$

where y_t is the observed dependent variable, $\boldsymbol{x}_t \in \Re^p$ and $z_t \in \Re^q$ are the vectors of regressors, $\boldsymbol{\beta}$ and $\boldsymbol{\delta}_j$ ($1 \le j \le m+1$) are the corresponding vectors of coefficients with $\boldsymbol{\delta}_i \neq \boldsymbol{\delta}_{i+1}$ ($1 \le i \le m$) and u_t is the error term. m is the number of structural breaks. The break dates $(T_1,\ldots,T_m)$ are explicitly treated as unknown and for $i=1,\ldots,m$, we have $\lambda_i = T_i/T$ with $0<\lambda_1<\cdots<\lambda_m<1$. Note that this is a partial structural change model in the sense that $\boldsymbol{\beta}$ is not subject to shift and is effectively estimated using the entire sample.[2] The purpose is to estimate the unknown regression coefficients and the break dates ($\boldsymbol{\beta}$, $\boldsymbol{\delta}_1,\ldots,\boldsymbol{\delta}_{m+1}$, $T_1,\ldots,T_m$) when T observations on (y_t, $\boldsymbol{x}_t$, z_t) are available. The above multiple linear regression model may be expressed in matrix form as

$$\mathbf{Y} = \mathbf{X}\beta + \bar{\mathbf{Z}}\delta + \mathbf{U},$$

where $\mathbf{Y}=(\boldsymbol{y}_1,\ldots,\boldsymbol{y}_T)'$, $\mathbf{X}=(\boldsymbol{x}_1,\ldots,x_T)'$, $\bar{\mathbf{Z}}$ is the matrix which diagonally partitions $\mathbf{Z}$ at the m-partition $(T_1,\ldots,T_m)$, i.e. $\bar{\mathbf{Z}}=\text{diag}(\mathbf{Z}_1,\ldots,\mathbf{Z}_{m+1})$ with $\mathbf{Z}_i=(z_{T_{i-1}}+1,\ldots,z_{T_i})'$, $\delta=(\boldsymbol{\delta}_1'$,

[2] When p=0, we obtain a pure structural change model where all the coefficients are subject to change.

$\boldsymbol{\delta}'_2,\ldots,\boldsymbol{\delta}'_{m+1})'$, and $\mathbf{U}=(\boldsymbol{u}_1,\ldots,\boldsymbol{u}_T)'$. Bai and Perron (1998) impose some restrictions on the possible values of the break dates. Indeed, they define the following set for some arbitrary small positive number ε: $\Lambda_\varepsilon=\{(\lambda_1,\ldots,\lambda_m);\ |\lambda_{i+1}-\lambda_i|\geq\varepsilon,\ \lambda_1\geq\varepsilon,\ \lambda_m\leq 1-\varepsilon\}$ to restrict each break date to be asymptotically distinct and bounded from the boundaries of the sample.

The estimation method considered is that based on the least-squares principle proposed by Bai and Perron (1998) . This method is described as follows. For each m-partition $(T_1,\ldots,T_m)$, denoted $\{T_j\}$, the associated least-squares estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\delta}_j$ are obtained by minimizing the sum of squared residuals $\sum_{i=1}^{m+1}\sum_{t=T_{i-1}+1}^{T_i}(y_t-\boldsymbol{x}'_t\boldsymbol{\beta}-\boldsymbol{z}'_t\boldsymbol{\delta}_i)^2$ where $T_0=0$ and $T_{m+1}=T$. Let $\hat{\beta}(\{T_j\})$ and $\hat{\delta}(\{T_j\})$ denote the resulting estimates. Substituting them in the objective function and denoting the resulting sum of squared residuals as $S_T(T_1,\ldots,T_m)$, the estimated break points $(\hat{T}_1,\ldots,\hat{T}_m)$ are such that

$$(\hat{T}_1,\ldots,\hat{T}_m)=\arg\min_{(T_1,\ldots,T_m)} S_T(T_1,\ldots,T_m), \tag{2}$$

where the minimization is taken over all partitions $(T_1,\ldots,T_m)$ such that $T_i-T_{i-1}\geq h$. Note that h is the minimal number of observations in each segment ($h\geq q$, not depending on T).[3] Thus, the break-point estimators are global minimizers of the objective function. Finally, the estimated regression parameters are the associated least-squares estimates at the estimated m-partition $\{\hat{T}_j\}$, i.e. $\hat{\beta}=\hat{\beta}(\{\hat{T}_j\})$ and $\hat{\delta}=\hat{\delta}(\{\hat{T}_j\})$. For our empirical applications, we use the efficient algorithm of Bai and Perron (2003a) based on the principle of dynamic programming which allows global minimizers to be obtained using a number of sums of squared residuals that is of order $O(T^2)$ for any $m\geq 2$.

3. Testing for structural change

Several tests for structural change have been proposed in the econometrics literature. These tests can be classified in two groups: (a) tests for a single structural change and (b) tests for multiple structural breaks.

3.1. The case of a single structural change

Chow (1960) considers a test for single structural change in which he imposes that the structural break must be known a priori and he uses a classic F statistic. The fact to consider a known break date a priori for this test implies that the investigator has only two choices: (a) to pick an arbitrary break date;[4] or (b) to pick a break date based on some known feature of the data series.[5] Consequently, the results can highly be sensitive to these arbitrary choices, and hence, the researchers can easily reach quite distinct conclusions. However, when the break date is unknown a priori, the problem

[3] From Bai and Perron (2003a), h should be set at $[\varepsilon T]$ if tests for structural change are used. But if the tests are not required, the estimation can be done with an arbitrary h greater than q.

[4] For this case, the Chow test may be uninformative, as the true break date can be missed.

[5] Here, the Chow test can be misleading, as the candidate break point is endogenous (it is correlated with the data) and the test is likely to falsely indicate a break when none in fact exists.

becomes complicated in the sense that this break date is a nuisance parameter that is present only under the alternative hypothesis of structural break. The standard asymptotic optimality properties of the Chow test do not then hold. Quandt (1960) resolves the problem by computing a sequence of Chow tests for each possible break date and estimates the break point as the date that maximizes the Chow test.[6] In the same context, Andrews (1993) proposes taking the largest Wald, Lagrange multiplier and likelihood ratio-like statistics over some possible break dates. His findings apply to a wide class of parametric models that are suitable for estimation by generalized method of moments (GMM) procedures. The break date may completely be unknown or it may be known to lie in a restricted interval. He suggests using the restricted interval [$0.15T$, $0.85T$] when no knowledge of the change point is available. He shows that the asymptotic distributions of these test statistics are nonstandard since, as we said above, the break-point parameter only appears under the alternative hypothesis. Andrews and Ploberger (1994) use a function that differs from the "sup" function and take exponential averages of the Wald, Lagrange multiplier and likelihood ratio-like tests sequences and show that improved power can be obtained. Andrews (1993) and Andrews and Ploberger (1994) tabulate critical values based on the derived nonstandard asymptotic distributions, and Hansen (1997) provides a method to calculate p-values. Hansen (2000) shows that these critical values are not robust to structural change in the marginal distribution of the regressors since the nonstandard asymptotic distributions of the test statistics are not the same; he then shows how to simulate robust critical values on a case-by-case basis.

The date which yields the largest value of the Wald test sequence is a good break date estimate if the estimated regression is linear and the Wald test is constructed with the homoskedastic form of the covariance matrix for the residuals[7]

$$\sup W_T = \sup_{T_1 \in [\pi T, (1-\pi)T]} W_T(T_1), \tag{3}$$

where $\pi \in (0,\ 0.5)$,

$$W_T(T_1) = \frac{\bar{S}_T - S_T(T_1)}{S_T(T_1)/(T-p-q)},$$

$\bar{S}_T$ is the sum of squared residuals under the null hypothesis and $S_T(T_1)$ is the sum of squared residuals under the alternative hypothesis, which depends on the break point T_1. However, an asymptotically equivalent version using the break date estimate $\hat{T}_1$ determined from the minimization of the sum of squared residuals can be obtained. The test is then

$$\sup W_T = W_T(\hat{T}_1),$$

[6] The test of Quandt (1960) is the likelihood ratio test under the hypothesis of normality.

[7] Bai (1997a) uses the Wald test in a sequential way to determine the number of breaks for the U.S. discount rate and the U.S. interest rate.

where the shift point estimate $\hat{T}_1$ is also obtained from the maximization of the $W_T(T_1)$ statistic.

When computing the Wald statistic using a heteroskedasticity-consistent covariance matrix, we obtain a different estimated break date; Table 1 reports estimates from this test and Table 2 reports those of the least-squares principle in cases where only a single break is allowed. The results demonstrate how these techniques can give different estimates of the break date for various time series. The difference is also revealed when using multiple break tests and estimation procedure instead of single break procedure. In this context, Bai (1997a), and Bai and Perron (2004) show that when multiple breaks are present, the power of tests for a single shift can be quite low in finite samples so that they might not lead us to reject the null hypothesis of no structural break when the true model has more than one break. Bai and Perron (1998) demonstrate how the least-squares estimator will converge to a global minimum coinciding with the dominating break in the presence of multiple structural changes.

The Wald test is significant for the break dates where the Wald test sequence lies above any critical value and vice versa. Hansen (2001) illustrates the usefulness of the Wald test computed using a heteroskedasticity-consistent covariance matrix through an assessment of the monthly U.S. labor productivity in the manufacturing/durables sector time series available from February 1947 to April 2001. He finds that the Wald test provides a break located in 1991:5.

3.2. The case of multiple structural breaks

A series of data can contain more than one structural break. To that effect, Bai and Perron (1998) recently provide a comprehensive analysis of several issues in the context of multiple structural change models and develop some tests which preclude the presence of trending regressors.

Table 1
Estimate results for the case of a single structural change

	Chow test[a]		Wald test[b]	
	Break date	Test statistic	Test statistic	Break date
Federal funds rate	1980:10	0.618	21.680	2000:7
	1981:6	5.654		
Federal discount rate	1977:11	1.010	20.853	1980:3
	1978:11	4.159		
Exchange rate euro/U.S. dollar	1993:8	0.555	25.760	1999:10
	1985:4	6.713		
Exchange rate yen/U.S. dollar	1980:12	0.478	17.826	1985:6
	1985:10	7.306		
Output	1919	0.104	14.377	1975
	1932	3.301		

[a] For the first date, the Chow test shows no evidence of a structural change, while for the second date, the test rejects the null hypothesis of no structural change at the 5% significance level.

[b] The Wald test is computed using a heteroskedasticity-consistent covariance matrix for the residuals.

Table 2
Estimate results for the case of a single structural change

	Wald test[a]		Least squares[b]	
	Test statistic	Break date	Break date	95% CI
Federal funds rate	13.818	1980:4	1980:4	(71:6–87:2)
Federal discount rate	38.782	1981:10	1981:10	(79:3–85:10)
Exchange rate euro/U.S. dollar	16.087	1985:3	1985:3	(81:12–87:4)
Exchange rate yen/U.S. dollar	26.196	1985:9	1985:9	(84:11–86:3)
Output	8.065	–	1933	(1921–1939)

[a] The Wald test is computed using the homoskedastic form of the covariance matrix for the residuals.
[b] In parentheses are the 95% confidence intervals for the estimated break date obtained using the least-squares principle.

3.2.1. A test of structural stability against a fixed number of breaks

Bai and Perron (1998) consider the sup F type test of no structural change (m=0) against the alternative hypothesis that there is a fixed number of breaks (m=k):

$$F_T(\lambda_1,\ldots,\lambda_k;q)=\frac{1}{T}\left(\frac{T-(k+1)q-p}{kq}\right)\frac{\hat{\delta}'\mathbf{R}'\left(\mathbf{R}\left(\bar{\mathbf{Z}}'\mathbf{M_X}\bar{\mathbf{Z}}\right)^{-1}\mathbf{R}'\right)^{-1}\mathbf{R}\hat{\delta}}{SSR_k},\tag{4}$$

where $\mathbf{R}$ is the conventional matrix such that $(\mathbf{R}\delta)'=(\delta_1'-\delta_2',\ldots,\delta_k'-\delta_{k+1}')$, $M_X=I-X(X'X)^{-1}X'$ and SSR_k is the sum of squared residuals under the alternative hypothesis, which depends on the break dates $(T_1,\ldots,T_k)$. The sup F type test statistic is defined as sup $F_T(k;q)=\sup_{(\lambda_1,\ldots,\lambda_k)\in\Lambda_\varepsilon}F_T(\lambda_1,\ldots,\lambda_k;q)$. However, we can obtain an asymptotically equivalent version using the break date estimates obtained from the global minimization of the sum of squared residuals. The test is then defined as

$$\sup F_T(k;q)=F_T\left(\hat{\lambda}_1,\ldots,\hat{\lambda}_k;q\right),$$

where the break fraction estimates $(\hat{\lambda}_1,\ldots,\hat{\lambda}_k)$ are also obtained from the maximization of the following F statistic since the break date estimators are consistent even in the presence of serial correlation:

$$F_T(\lambda_1,\ldots,\lambda_k;q)=\frac{1}{T}\left(\frac{T-(k+1)q-p}{kq}\right)\hat{\delta}'\mathbf{R}'\left(\mathbf{R}\hat{\mathbf{V}}(\hat{\delta})\mathbf{R}'\right)^{-1}\mathbf{R}\hat{\delta},$$

where $\hat{\mathbf{V}}(\hat{\delta})=(\bar{Z}'M_X\bar{Z}/T)^{-1}$ is the covariance matrix of $\hat{\delta}$ assuming spherical errors. This test is a generalization of the sup F test considered by Andrews (1993) and others for the case k=1. Different versions of these tests can be obtained depending on the assumptions made with respect to the distribution of the regressors and the errors across segments (e.g., Bai and Perron, 2003a, 2004).

3.2.2. A test of structural stability against an unknown number of breaks

They also consider tests of no structural change against an unknown number of breaks given some upper bound M for m. The following new class of tests is called double maximum tests and is defined for some fixed weights $\{a_1,\ldots,a_M\}$ as

$$D\max F_T(M,q,a_1,\ldots,a_M) = \max_{1\leq m\leq M} a_m \sup_{(\lambda_1,\ldots\lambda_m)\in\Lambda_\varepsilon} F_T(\lambda_1,\ldots,\lambda_m;q). \tag{5}$$

We can use the asymptotically equivalent version:

$$D\max F_T(M,q,a_1,\ldots,a_M) = \max_{1\leq m\leq M} a_m F_T(\hat{\lambda}_1,\ldots\hat{\lambda}_m;q).$$

The weights $\{a_1,\ldots,a_M\}$ reflect the imposition of some priors on the likelihood of various numbers of structural breaks. In a first time, Bai and Perron (1998) set all weights equal to unity (a_m=1) and label this version of the test as *UD* max F_T (M, q). They also consider a set of weights such that the marginal p-values are equal across values of m. The weights are then defined as a_1=1 and for m>1 as $a_m=c(q,\ \alpha,\ 1)/c(q,\ \alpha,\ m)$, where α is the significance level of the test and $c(q,\alpha,m)$ is the asymptotic critical value of the test $\sup_{(\lambda_1,\ldots,\lambda_m)\in\Lambda_\varepsilon} F_T(\lambda_1,\ldots,\lambda_m;\ q)$. This version of the test is denoted as *WD* max $F_T(M,\ q)$.

Note that the significance of the sup $F_T(k;\,q)$, the *UD* max $F_T(M,\,q)$ and the *WD* max $F_T(M,\,q)$ tests does not provide enough information about the exact number of breaks, but only indicates that one break is at least present.

3.2.3. A test of l against (l+1) breaks

The last test developed by Bai and Perron (1998) is a sequential test of l versus (l+1) structural breaks:

$$\sup\ F_T(l+1|l) = \{S_T(\hat{T}_1,\ldots,\hat{T}_l) - \min_{1\leq i\leq l+1} - \inf_{\tau\in\Lambda_{i,\varepsilon}} S_T(\hat{T}_1,\ldots,\hat{T}_{i-1},\tau,\hat{T}_i,\ldots,\hat{T}_l)\}/\hat{\sigma}^2, \tag{6}$$

where $\Lambda_{i,\varepsilon}=\{\tau;\ \hat{T}_{i-1}+(\hat{T}_i-\hat{T}_{i-1})\varepsilon\leq\tau\leq\hat{T}_i-(\hat{T}_i-\hat{T}_{i-1})\varepsilon\}$, $S_T(\hat{T}_1,\ldots,\hat{T}_{i-1},\tau,\ \hat{T}_i,\ldots,\hat{T}_l)$ is the sum of squared residuals resulting from the least-squares estimation from each m-partition $(T_1,\ldots,T_m)$ and $\hat{\sigma}^2$ is a consistent estimator of σ^2 under the null hypothesis. To compute the test statistic, the method consists in applying (l+1) tests of the null hypothesis of stability against the alternative hypothesis of a single break. The test is applied to each segment $[\hat{T}_{i-1}+1,\ \hat{T}_i]$ for $i=1,\ldots,l+1$, and with $\hat{T}_0=0$ and $\hat{T}_{l+1}=T$.[8] We reject the null hypothesis and we conclude in favor of a model with (l+1) structural breaks if the sum of squared residuals obtained from the estimated model with l changes is sufficiently larger than the overall minimal value of the sum of squared residuals (over all segments where an additional change is included), and the break point thus selected is the one associated with

[8] Given the estimates $\hat{T}_i$, we require that the break fractions $\hat{\lambda}_i=\hat{T}_i/T$ converge to their true values at rate T. Hence, the estimates $\hat{T}_i$ need not be obtained by a global minimization of the sum of squared residuals; we can also use the sequential one-at-a-time estimates which imply break fractions that converge at rate T (e.g., Bai, 1997b).

this overall minimum.[9] This test is used in a sequential way to estimate consistently the number of changes in a set of data (see Section 4.1).

The asymptotic distributions of all these tests are derived in Bai and Perron (1998) and asymptotic critical values are tabulated in Bai and Perron (1998, 2003b) for ε=0.05, 0.10, 0.15, 0.20 and 0.25. The maximum permitted number of changes M is 9 for ε=0.05, 8 for ε=0.10, 5 for ε=0.15, 3 for ε=0.20 and 2 for ε=0.25. Bai and Perron (2003a) present a few empirical applications to illustrate the usefulness of these tests. Indeed, they analyze the U.S. ex-post real interest rate series considered by Garcia and Perron (1996) and reevaluate some findings of Alogoskoufis and Smith (1991),[10] who analyze structural changes in the persistence of inflation and shifts in an expectations-augmented Phillips curve resulting from such changes using postwar annual data from the United Kingdom and the United States. They find that some tests are significant.

Bai (1999) considers a likelihood ratio test allowing for the presence of trending regressors. The test statistic is based on the difference between the optimal sum of squared residuals associated with l breaks and the optimal sum of squared residuals associated with (l+1) breaks as the sup $F_T(l+1|l)$ test. The test statistic is denoted as

$$\sup\ LR_T(l+1|l) = \frac{S_T(\hat{T}_i,\ldots,\hat{T}_l) - S_T(\hat{T}_1,\ldots,\hat{T}_{l+1})}{S_T(\hat{T}_1,\ldots,\hat{T}_{l+1})/T}, \tag{7}$$

He shows that asymptotic critical values can be asymptotically obtained since the limiting distribution of the test statistic has a known analytical density function. Note that when the errors are i.i.d. normal random variables, the test is called a likelihood ratio test. On the other hand, when the normality is not assumed, the test may be considered as a pseudo-likelihood ratio test. As the last test of Bai and Perron (1998), this test is also used in a sequential way to estimate the number of structural breaks.

Some tests are proposed in the context of co-integrated regression models by Hansen (1992a) who derives the large-sample distributions of Lagrange multiplier tests for parameter instability against several alternatives. He includes the Sup$F$ test of Quandt (1960), as well as the Lagrange multiplier tests of Nyblom (1989) and Nabeya and Tanaka (1988). He applies the tests to the U.S. aggregate consumption function, the present-value model of U.S. stock prices and dividends[11] and the term structure of U.S. interest rates. He finds that some relationships are not stable and that the change in the Federal Reserve's operating procedures altered the relationship between some interest rates. He shows that this regime shift only appears to have affected the relationship between the federal funds

[9] In other words, we conclude in favor of a model with (l+1) changes if the overall maximal value of the sup $F_T(1;\, q)$ (over all segments where an additional break is included) is sufficiently large and the break date thus chosen is the one associated with this overall maximum.

[10] Alogoskoufis and Smith (1991) analyze structural changes in the persistence of inflation and shifts in an expectations-augmented Phillips curve resulting from such changes using postwar annual data from the United Kingdom and the United States. They argue that the process which describes inflation exhibits a structural change from 1967 to 1968 for the United Kingdom time series.

[11] Campbell and Shiller (1987) argue that a standard rational-expectations model of asset markets implies that real stock prices and dividends should be co-integrated. Campbell and Shiller (1988a,b) argue for a logarithmic approximation that implicitly assumes that the logarithms of the price and dividend indexes are co-integrated. The data used are annual and cover the period 1871–1986.

rate and the treasury-bill rates but not the relationship between the treasury-bill rates of different maturities.

Once we are sure that the series are characterized by the presence of structural changes using the sup $F_T(k; q)$, the UD max $F_T(M, q)$ and the WD max $F_T(M, q)$ tests, we are in the stage of determination of the number of breaks using some selection procedures.

4. Estimating the number of breaks

For the study and the analysis of structural change models, the estimation of the number of structural breaks receives an important attention from the researchers. Several selection procedures exist in the literature, including the sequential procedures based on a sequence of tests and the information criteria.

4.1. The sequential procedures

Bai and Perron (1998) suggest a method based on a sequential application of the sup $F_T(l+1|l)$ test.[12] In a first time, we briefly recall the sequential method one-at-a-time when the number of breaks is known, say n. Once the first break date is identified, the sample is split into two subsamples separated by this first estimated break date. For each subsample, we estimate a model with one break and the second break date is chosen as that break date which allows the greatest reduction in the sum of squared residuals. The sample is then partitioned in three subsamples and a third break date is chosen as the estimate from three estimated one-break models that allows the greatest reduction in the sum of squared residuals. This process is continued until the n break dates are selected. Note that the obtained estimates are not guaranteed to be identical to those obtained by global minimization of the sum of squared residuals.

The sequential procedure to estimate the number of changes is now the following:

- We start by estimating a model with a small number of structural breaks,[13] using the global minimization of the sum of squared residuals or the sequential method one-at-a-time since both imply break fractions that converge at rate T (e.g., Bai, 1997b).
- We then perform parameter constancy tests for each subsample (those obtained by cutting off at the estimated break points), adding a break to a subsample associated with a rejection with the test sup $F_T(l+1|l)$.
- This process is repeated by increasing l sequentially until the test sup $F_T(l+1|l)$ fails to reject the no additional structural change hypothesis.

The final number of break points is thus equal to the number of rejections obtained with the parameter constancy tests plus the number of changes used in the initial step. Moreover, this method allows us to obtain the break-point estimators.

[12] For the application of this sequential method, the readers are referred to Bai and Perron (2003a, 2004).

[13] We can start by estimating a model with no break.

Bai and Perron (2004, 2003a) show that we can improve the performance of the sequential procedure in the context of estimating the number of breaks. Indeed, a preferred strategy is to first look at the *UD* max or *WD* max tests to see if at least a structural break exists. We can then decide the number of breaks based upon an examination of the sup $F_T(l+1|l)$ statistics constructed using the break date estimates obtained from a global minimization of the sum of squared residuals [i.e. we select m breaks such that the tests sup $F_T(l+1|l)$ are nonsignificant for any $1 \geq m$]. Bai and Perron (2004) use this strategy to determine the number of breaks based on simulated data. Bai and Perron (2003a) conclude that this method leads to the best results and is recommended for empirical applications. They also illustrate its usefulness to determine the number of structural changes for the postwar annual U.K. inflation rate series available from 1947 to 1987. They find that the sequential procedure selects zero break but the significance of the sup $F_T(2)$, the *UD* max, the *WD* max and the sup $F_T(2|1)$ tests suggests a model with two breaks estimated at 1967 and 1975. The first break date is linked to the end of the Bretton Woods system. Another sequential procedure based on the application of the sup $LR_T(l+1|l)$ test is considered by Bai (1999) to estimate consistently the number of breaks in the data.[14]

4.2. The model selection criteria

The other procedures consist in estimating the number of structural breaks using the information criteria. The basic idea is that one must penalize the addition of any break point since the sum of squared residuals is monotonically decreasing in m. The penalty must force the estimator of the number of break points to converge rapidly to the true value to assure the asymptotic properties of all the estimations that depend on that estimator. For easily identified models, a large penalty greatly reduces the probability of overestimation of the number of breaks. However, if the model is difficult to identify (for example, if $\|\delta_{j+1}-\delta_j\|$ is small), then a large penalty results in probable underestimation of the number of changes.

We recall that the Schwarz (1978) criterion is defined as

$$SIC(m) = \ln\left(S_T\left(\hat{T}_1,\ldots,\hat{T}_m\right)/(T-m)\right) + 2p^*\ln(T)/T, \tag{8}$$

where $p^*=(m+1)q+m+p$ is the number of unknown parameters. Yao (1988) suggests the use of the following Bayesian information criterion:

$$BIC(m) = \ln\left(S_T\left(\hat{T}_1,\ldots,\hat{T}_m\right)/T\right) + p^*\ln(T)/T. \tag{9}$$

He shows that the estimator of the number of breaks $\hat{m}$ converges (at least for normal sequence of random variables with shifts in mean) to m^0, the true number of breaks, provided $m^0 \leq M$. Yao and Au (1989) propose the criterion

$$YIC(m) = \ln\left(S_T\left(\hat{T}_1,\ldots,\hat{T}_m\right)/T\right) + mC_T/T, \tag{10}$$

where $\{C_T\}$ is any sequence satisfying $C_T T^{-2/v} \to \infty$ and $C_T/T \to 0$ as $T \to \infty$ for some positive integer v. The error term is with finite 2vth moment for any $v \geq 3$. In our empirical

[14] This procedure is not presented and is beyond the scope of the paper.

applications, we use the sequence $C_T = 0.368T^{0.7}$ proposed by Liu et al. (1997) who propose the following modified Schwarz' criterion:

$$MIC(m) = \ln\left(S_T\left(\hat{T}_1,\ldots,\hat{T}_m\right)/(T-p^*)\right) + p^* c_0 [\ln(T)]^{2+\delta_0}/T. \tag{11}$$

They suggest using c_0=0.299 and δ_0=0.1 based on the performance of the estimator of the number of changes obtained by the criterion MIC for various simulation experiments carried out with some models. The estimated number of break points $\hat{m}$ is obtained by minimizing the above-mentioned information criteria given an upper bound M for m.

Nunes et al. (1996) show that the criterion BIC tends to select the maximum possible number of changes for an integrated process of order one without breaks when we estimate a model with change in mean and change in trend. Perron (1997) studies via simulations the behavior of the information criteria BIC and MIC in the context of estimating the number of breaks in the trend function of a series in the presence of serial correlation. These criteria perform reasonably well when the errors are uncorrelated but choose a number of changes much higher than the true value when serial correlation is present. When the errors are uncorrelated but a lagged dependent variable is present, the criterion BIC performs badly when the coefficient on the lagged dependent variable is large (and more so as it approaches unity). In such cases, the criterion MIC performs better under the null hypothesis of no break but underestimates the number of structural breaks when some are present. The results of Perron (1997) show that the conclusions of Nunes et al. (1996) do not depend on the fact that the data-generating process is a random walk; even an AR(1) process with a correlation degree smaller than 1 leads to an overestimation of the number of breaks. In the same context, Boutahar and Jouini (2003a) prove that when the data-generating process is a trend-stationary process or stationary autoregressive process without any structural change, the above-mentioned criteria have a tendency to overestimate the number of breaks when we run a regression with change in mean.

The distinct advantage of the sequential method of Bai and Perron (1998) is that, unlike the information criteria, it can directly take into account the presence of serial correlation in the errors and heterogeneous variances across segments. Bai and Perron (2003a, 2004) favor the sequential method based on the sup $F_T(l+1|l)$ test which seems to perform better than procedures based on information criteria.

After the determination of the number of breaks by the different procedures, we attempt to estimate the locations of the breaks, their confidence intervals and the regression parameters of the regimes.

5. Estimating the locations of breaks

Another important aspect in the study of structural change models is to know when the structural break occurs. In linear regression models, the appropriate estimation method is that based on the least-squares principle (see Section 2). Bai (1994,1997a) studies the case of a single structural break. Indeed, he derives the asymptotic distribution of the break-point estimator allowing the construction of confidence intervals which indicate the degree of estimation accuracy. Chong (2001) develops a comprehensive asymptotic theory for an autoregressive process of order one [AR(1)] with an unknown single structural break. More precisely, he examines the case where an AR(1) process changes from a stationary one to a nonstationary one (or the other way around). In each case, he establishes the consistency of the estimators and derives their limiting distributions.

Bai and Perron (1998) extend the analysis to multiple structural change models. They establish the limiting distribution of the break date estimators for shifts with shrinking magnitudes. We can then construct confidence intervals for the break date estimates under various assumptions on the structure of the regressors and the errors in the different segments (e.g., Bai and Perron, 2003a, 2004). In the same context, Bai et al. (1998) show that using multiple time series improves the estimation precision. They attempt to date the alleged slowdown of the early 1970s and show that there is no evidence of structural change when examining the individual series with univariate models, but there is strong evidence in a joint vector autoregression. They use the U.S. quarterly data for 1959 through 1995 on real output, consumption and investment and find a structural change at the first quarter of 1969 and their 90% confidence interval is (1966:2–1971:4). The estimation of multiple break dates can be done using a sequential estimation method (e.g., Bai, 1997b; Chong, 1995). The idea is that the sum of squared residuals—as a function of break date—can have a local minima near each break date when there are multiple structural changes in the process. The global minimum can thus be used as a break-point estimator, while the other local minima can be viewed as candidate break-point estimators. The sample is then partitioned at the estimated break date, and the analysis continues on the subsamples. The advantage of this method lies in its robustness to misspecification in the number of changes and its computational savings. The number of least-squares regressions required to compute all the break dates is of order T, the sample size. Each estimated break point is consistent for one of the true ones despite underspecification of the number of structural changes. Bai (1997b) shows that important improvements are obtained by a re-estimation of these sequential break points based on refined samples; this procedure is as follows. If a sample or subsample contains more than one break, then an obtained estimated change point should be re-estimated. For example, suppose that there are two breaks $\hat{T}_1$ and $\hat{T}_2$ in a series. If $\hat{T}_1$ is identified in the first place, i.e. from the whole sample, then $\hat{T}_1$ should be re-estimated using the subsample $[1, \hat{T}_2]$. On the other hand, if $\hat{T}_2$ is identified in advance of $\hat{T}_1$, then $\hat{T}_2$ should be re-estimated using the subsample $[\hat{T}_1, T]$. Note that each refined estimator has the same asymptotic distribution as if the sample had a single break date. The refinement may be performed prior to the complete estimation of all the structural breaks. Chong (1994) provides a comprehensive analysis of underspecifying the number of break dates in structural change models. More precisely, he allows for stochastic regressors and time trend in the model. Chong (1995) shows that the break fraction estimator in a single structural change regression applied to data that contain two mean shifts converges to one of the two true break fractions, namely, the one which allows the greatest reduction in the sum of squared residuals.

We now discuss a repartition procedure that yields an estimator having the same asymptotic distribution as the simultaneous estimators.[15] Once initial T-consistent

[15] The simultaneous estimators are those obtained using the global minimization of the sum of squared residuals as we have showed in Section 2.

estimators are obtained, the repartition technique re-estimates each of the break dates based on the initial estimates. For example, suppose that we have two initial T-consistent estimators $\hat{T}_1$ and $\hat{T}_2$. To estimate T_1, we use the subsample $[1, \hat{T}_2]$, and to estimate T_2, the subsample $[\hat{T}_1, T]$ is used. The new estimators are also T-consistent for the true parameters and have limiting distributions identical to what it would be for a single break-point model or for a model with multiple breaks estimated by the simultaneous method described in Section 2.

The estimates of unknown shift points are particularly sensitive to the sample period. Indeed, the estimated break dates that we obtain might be different from those obtained by other authors using the same data but not the same period. This is revealed by the fact that the objective function of the estimation method depends on the break parameters. Consequently, possible break dates that were ignored in the smaller sample can be taken into account when the data cover a more larger sample. This can be viewed in Ben Aïssa and Jouini (2003) and Boutahar and Jouini (2003b) who show that using two monthly U.S. inflation series which do not cover the same period and based on some information criteria, there is an additional break date detected for the larger sample.

6. Empirical illustrations

Using the postwar monthly U.S. inflation rate (seasonally adjusted) covering the period 1956:1–2002:9, Ben Aïssa and Jouini (2003) show that using some model selection criteria, the evolution curve of inflation in the United States was flattened during the last 20 years since it is noted that this reduction in extent of inflation is stable and durable. Unlike Ben Aïssa and Jouini (2003), Jouini and Boutahar (2003) use the same series to show that using the Bai and Perron (1998) sequential selection procedure, the U.S. inflation process is unstable after 1982:6 as there is a break at the beginning of the 1990s. In the same context, Ben Aïssa et al. (2004) use a test similar to the one based on Kolmogorov–Smirnov statistic applied to the evolutionary spectrum, and Bai and Perron (1998) procedure based on the UD max $F_T(M, q)$, WD max $F_T(M, q)$ and sup $F_T(l+1|l)$ tests to estimate the number of breaks and their locations in the monthly U.S. inflation series covering the period 1957:1–2003:4. The obtained results of the two approaches are similar and economically significant. They also find that inflation was perfectly stable and durable during the 1990s. However, at the beginning of the 2000s, the U.S. economy was marked by a light recession expressed by a decrease of productivity and an increase in unemployment and inflation.

We now apply some of the procedures presented in this paper to U.S. time series, namely, the interest rates—the federal funds rate and the federal discount rate—the exchange rates euro/US dollar and yen/US dollar and the gross national product.[16] We illustrate the usefulness of some tests for the case of single and multiple structural changes, selection procedures to estimate the number of breaks and methods to form confidence intervals for the break dates. To that effect, we adopt an AR(1) process with drift to

[16] See Appendix A for the definition and sources of the data.

describe the time series, and our approach is directly oriented at the issue of testing for structural changes in the level and the persistence of the series, i.e. $\boldsymbol{x}_t=0$ and $z_t=(1, y_{t-1})'$.

6.1. The case of a single break

We are interested in testing the presence of structural changes using the Chow and the Wald tests and estimating the break date using the least-squares principle.

- For the Chow test, we compute the test statistic for two different breaks arbitrarily picked to show that, at the 5% significance level, the results are highly sensitive to these arbitrary choices and hence we can easily reach distinct conclusions. For the first date, the Chow test shows no evidence of a structural change, while for the second date, the test rejects the null hypothesis of no structural break.
- For the Wald test and the least-squares principle, we compute the test statistic and we estimate the shift point over all candidate break dates contained in the restricted interval $[0.05T, 0.95T]$ since we cannot consider break dates too close to the boundaries of the sample, as there are not enough observations to identify all the subsample parameters.
- For the Wald test, we compute two statistics constructed with the homoskedastic form of the covariance matrix and using a heteroskedasticity-consistent estimator of this matrix.
- The Andrews (1993) critical values are 16.44 (1%), 12.93 (5%) and 11.20 (10%).

From the results available in Tables 1 and 2, we observe that the estimated break date obtained using the Wald test constructed with the homoskedastic form of the covariance matrix and the one obtained by minimizing the sum of squared residuals are identical and differ from the one obtained using the Wald test computed using a heteroskedasticity-consistent covariance matrix. The Wald test constructed assuming the homoskedasticity and the heteroskedasticity is significant since the sequences of Wald statistics cross the Andrews 5% critical value several times, achieving maximal value at a break date (see graphs in Appendix B).[17] We are now confident that the time series have at least one structural break. Except for the exchange rate yen/U.S. dollar, the break date for the other series is imprecisely estimated since the 95% confidence interval covers a large period.

Using 10% trimming (the results are not reported, but available from the authors), the Wald test constructed using a heteroskedasticity-consistent covariance matrix gives a shift point estimator different from the one obtained using 5% trimming for the federal funds rate (1989:5) and the exchange rate euro/U.S. dollar (1985:3). This confirms our noticing that the results are particularly sensitive to the sample period used for the estimation of the break date. On the other hand, the variation of the trimming does not affect the break date estimate obtained by the least-squares principle for all the time series. This might be explained by the fact that the shift point is strongly evident. As we will show in the next

[17] For the output, the Wald test constructed using the homoskedastic form of the covariance matrix is not significant since the maximal value of the sequence of Wald statistics lies below the Andrews 5% critical value (see graph in Appendix B).

section, this estimated break date is one of the estimated break points obtained using the estimation procedure of the case of multiple breaks.

6.2. The case of multiple changes

We now test for structural changes using the procedures developed in Bai and Perron (1998); more precisely, we estimate the number of structural breaks as well as their locations.

- We impose the minimum structure on the data by allowing for different distributions of both the regressors and the errors in the different segments. Indeed, the majority of the graphs of the series (Appendix B) show different variability in different periods. Hence, we investigate the stability of the process allowing for different variances for the residuals across segments.[18]
- No serial correlation is permitted in the errors $\{u_t\}$ since we consider the case where a lagged dependent variable is allowed as regressor [see Part (ii) of Assumption A4 in Bai and Perron, 1998].
- When constructing confidence intervals for the break dates, we assume that the distributions of the regressors are heterogeneous across subsamples.
- The maximum permitted number of breaks is set at M=5, and we use a trimming ε=0.10 which means that the minimal number of observations in each segment is set at $h=[\varepsilon T]$, with T the sample size.
- Throughout the remainder of this paper, sup $F_T(k)$, UD max and WD max denote, respectively, the sup $F_T(k;\, q)$, UD max $F_T(M,\, q)$ and WD max $F_T(M,\, q)$ tests.
- At 5% significance level and for ε=0.10, the critical values for the sup $F_T(k)$ tests for k=1,. . .,5 are, respectively, 12.25, 10.58, 9.29, 8.37 and 7.62, for the double maximum tests UD max and WD max they are, respectively, 12.59 and 13.66 and for the sup $F_T(l+1|l)$ tests for l=1,. . .,4 they are, respectively, 13.83, 14.73, 15.46 and 16.13.
- The repartition procedure re-estimates each of the break dates based on the estimates obtained using the sequential method (Section 4.1) since they are T-consistent estimators.

6.2.1. The U.S. interest rates

The results of the different procedures are presented in Tables 3 and 4. The first issue to consider is testing for structural changes. The sup $F_T(k)$ tests are all significant at the 5% level for k=1,. . .,5 except for the federal funds rate where it is not the case for k=2 even at the 10% significance level.[19] The double maximum tests UD max and WD max which allow us to test the null hypothesis of no structural break versus an unknown number of changes given the upper bound of five breaks are significant at the 5% level. The significance of these tests does not provide enough information about the exact number of breaks but means that one break is at least present. We now turn to the sup $F_T(l+1|l)$ tests

[18] Note that the existence of breaks in the variance could be exploited to increase the precision of the break date estimates (e.g., Bai and Perron, 2003a).

[19] For the sup $F_T(2)$ test, the critical value at the 10% significance level is 9.43.

Table 3
Estimate results for the federal funds rate (1957:1–2002:10)

		ε=0.10	M=5	T=550		
Tests	sup F_T (1)	sup F_T (2)	sup F_T (3)	sup F_T (4)	sup F_T (5)	UD max
	15.805*	8.831**	10.222*	9.056*	9.796*	15.805*
	WD max	sup F_T (2|1)	sup F_T (3|2)	sup F_T (4|3)	sup F_T (5|4)	
	15.805*	17.538*	17.538*	5.117	6.085	
	SIC	BIC	YIC	MIC	Seq	Rep
Number of breaks[a]	0	0	0	0	2	2
Estimates with three breaks[b]	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$			
	1975:10 (73:10–79:2)	1980:4 (77:11–80:9)	1984:10 (84:9–84:11)			
	$\hat{\delta}_{1,1}$	$\hat{\delta}_{2,1}$	$\hat{\delta}_{1,2}$	$\hat{\delta}_{2,2}$	$\hat{\delta}_{1,3}$	$\hat{\delta}_{2,3}$
	0.087 (0.060)	0.984 (0.011)	−0.333 (0.189)	1.071 (0.022)	1.309 (0.737)	0.881 (0.058)
	$\hat{\delta}_{1,4}$	$\hat{\delta}_{2,4}$				
	0.013 (0.051)	0.990 (0.008)				

[a] Seq and Rep mean, respectively, the sequential and the repartition procedures. We use a 5% significance level for these procedures.
[b] In parentheses, reported are the standard errors (robust to serial correlation) for the estimated regression coefficients and the 95% confidence intervals for the break dates.
* Indicates that the tests are significant at the 5% level.
** Indicates that the sup $F_T(2)$ test is not significant even at the 10% level (the corresponding critical value is 9.43).

which are nonsignificant for any $l \geq 3$ at the 5% level. The sequential procedure using a 5% level selects two breaks for the two series and the information criteria detect zero breaks except for the federal discount rate where the criteria BIC and YIC choose one break located in 1981:10. As we mentioned in Section 4.1, based on the results of the tests, a preferred strategy to determine the number of changes is to first look at the *UD* max or *WD* max tests to see if at least a break is present and the number of structural breaks can be decided based upon a sequential examination of the sup $F_T(l+1|l)$ statistics constructed using global minimizers for the break dates. Indeed, the significance of the *UD* max, the *WD* max and the sup $F_T(3|2)$ tests and the nonsignificance of the sup $F_T(l+1|l)$ test for any $l \geq 3$ suggest a model with three breaks. Given the obtained break dates (see Tables 3 and 4), it is interesting to note that the estimated break dates are remarkably similar for the two interest rates. Another feature of substantial importance is that the first two break dates coincide with the second and third breaks detected in the U.S. inflation rate (1973:9 and 1982:6) by Jouini and Boutahar (2003). We would have to consider the possibility that there is a relationship between nominal interest rates and inflation. Note that since the late 1980s, there has been a large increase in the application of co-integration techniques to uncover this relationship. We can deduce that the instability of the series is due to the 1973 Oil-Price Shock, the revision of IMF's Charter on January 1976, the change in the Federal Reserve's operating procedures in 1979, the second Oil-Price Shock and the Plaza Accord on September 1985. The third break date for the federal funds rate and the second break date for the federal discount rate have small 95% confidence intervals indicating that the breaks are precisely estimated. The other break dates are, however, imprecisely estimated since their 95% confidence intervals cover a large period. The differences in the estimated coefficients on the lagged dependent variable over each segment are insignificant since they point to slight variations over segments.

6.2.2. Exchange rate dynamics

The corresponding results are reported in Tables 5 and 6. The sup $F_T(k)$ (k=1,. . .,5), the *UD* max and the *WD* max tests are all significant at the 5% significance level indicating the presence of at least one structural change in the data. The sup $F_T(l+1|l)$ test is not significant at the 5% level for any $l \geq 2$ for the exchange rate euro/U.S. dollar and for any $l \geq 1$ for the exchange rate yen/U.S. dollar, indicating the presence of two and one break, respectively. The sequential procedure using a 5% significance level selects two breaks for the first rate and one for the second. On the other hand, the information criteria underestimate the number of breaks since they choose zero break except for the criteria BIC and YIC which select one break located in 1985:9 for the exchange rate yen/U.S. dollar. Given the fact that the procedures based on information criteria are biased downward, we favor the sequential procedure and the sup $F_T(l+1|l)$ test which perform better in this case and we conclude in favor of a model with two changes estimated at 1985:3 and 1999:1 for the exchange rate euro/U.S. dollar and a model with one break located in 1985:9 for the exchange rate yen/U.S. dollar.[20] The break date 1985:3 has rather large 95% confidence interval (between 1982:12 and 1986:3), while the break dates 1985:9 and 1999:1 are, however, precisely estimated since their 95% confidence intervals

[20] These break date estimates are obtained under global minimization.

Table 4
Estimate results for the federal discount rate (1957:1–2002:10)

		ε=0.10	M=5	T=550		
Tests	sup F_T (1)	sup F_T (2)	sup F_T (3)	sup F_T (4)	sup F_T (5)	UD max
	39.890*	12.795*	12.621*	14.337*	12.418*	39.890*
	WD max	sup F_T (2|1)	sup F_T (3|2)	sup F_T (4|3)	sup F_T (5|4)	
	39.890*	15.911*	15.911*	6.650	4.212	
	SIC	BIC	YIC	MIC	Seq	Rep
Number of breaks	0	1	1	0	2	2
Estimates with three breaks	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$			
	1977:4 (72:1–77:7)	1981:10 (81:3–83:2)	1986:4 (85:10–88:2)			
	$\hat{\delta}_{1,1}$	$\hat{\delta}_{2,1}$	$\hat{\delta}_{1,2}$	$\hat{\delta}_{2,2}$	$\hat{\delta}_{1,3}$	$\hat{\delta}_{2,3}$
	0.043 (0.032)	0.992 (0.006)	0.312 (0.210)	0.984 (0.020)	0.583 (0.186)	0.921 (0.020)
	$\hat{\delta}_{1,4}$	$\hat{\delta}_{2,4}$				
	−0.045 (0.042)	1.003 (0.008)				

* Indicates that the tests are significant at the 5% level.

Table 5
Estimate results for the exchange rate euro/U.S. dollar (1980:1–2001:5)

		ε=0.10	M=5	T=257		
Tests	sup F_T (1)	sup F_T (2)	sup F_T (3)	sup F_T (4)	sup F_T (5)	UD max
	17.161*	17.137*	14.533*	11.246*	10.368*	17.161*
	WD max	sup F_T $(2\mid 1)$	sup F_T $(3\mid 2)$	sup F_T $(4\mid 3)$	sup F_T $(5\mid 4)$	
	19.842*	20.123*	10.981	4.280	6.972	
	SIC	BIC	YIC	MIC	Seq	Rep
Number of breaks	0	0	0	0	2	2
Estimates with two breaks	$\hat{T}_1$	$\hat{T}_2$				
	1985:3 (82:12–86:3)	1999:1 (98:3–99:11)				
	$\hat{\delta}_{1,1}$	$\hat{\delta}_{2,1}$	$\hat{\delta}_{1,2}$	$\hat{\delta}_{2,2}$	$\hat{\delta}_{1,3}$	$\hat{\delta}_{2,3}$
	0.006 (0.015)	0.981 (0.015)	0.073 (0.018)	0.939 (0.016)	0.087 (0.050)	0.900 (0.051)

* Indicates that the tests are significant at the 5% level.

Table 6
Estimate results for the exchange rate yen/U.S. dollar (1971:1–2003:4)

		ε=0.10	M=5	T=388		
Tests	sup F_T (1)	sup F_T (2)	sup F_T (3)	sup F_T (4)	sup F_T (5)	UD max
	38.980*	24.801*	21.834*	17.931*	15.267*	38.980*
	WD max	sup F_T (2|1)	sup F_T (3|2)	sup F_T (4|3)	sup F_T (5|4)	
	38.980*	13.480	13.480	13.480	15.137	
	SIC	BIC	YIC	MIC	Seq	Rep
Number of breaks	0	1	1	0	1	1
Estimates with one break	$\hat{T}_1$					
	1985:9 (84:11–86:2)					
	$\hat{\delta}_{1,1}$	$\hat{\delta}_{2,1}$	$\hat{\delta}_{1,2}$	$\hat{\delta}_{2,2}$		
	6.655 (3.121)	0.971 (0.011)	8.844 (1.463)	0.925 (0.011)		

* Indicates that the tests are significant at the 5% level.

cover only a few months before and after. The autoregressive coefficients have nearly unit roots and slightly decrease over segments. The fact to find two break dates during the year 1985 for the two rates reinforces the existence of a break for the American dollar during this period. Note that the date 1985:9 can be linked to the Plaza Accord on September 1985.

6.2.3. The U.S. output

We observe, from the graph of the Gross National Product (Appendix B), that the series is with values increasing over the sample period and that it may be affected by structural breaks occurring during the period 1929–1945. The corresponding results are given in Table 7. The majority of tests are not significant even at the 10% level like the sup $F_T(l+1|l)$ tests for all l. The information criterion YIC detects three breaks, while the others and the sequential procedure choose zero change. As the other series, we determine the number of breaks based on the results of the tests. Indeed, the significance of the *WD* max test at the 10% level[21] allows us to conclude that there is at least one break in the series. We then choose a model with one break estimated at 1933 since the sup $F_T(l+1|l)$ test is nonsignificant for any $l\geq 1$. This reinforces our noticing mentioned above about the break date. This date can be linked to the Great Depression period, 1930–1933. Note that the 95% confidence interval is large indicating that the break is rather imprecisely estimated. We remark that the estimated autoregressive coefficient does not almost vary across segments.

6.2.4. Comments

- We first remark that the estimated autoregressive coefficient slightly varies across segments for all the series explaining without doubt the inadequate results of the information criteria, especially the SIC and MIC, which are biased downward since they underestimate the number of structural changes when the model is difficult to identify. To confirm this, we have carried out simulation experiments (not reported here, but available upon request from the authors) in which the data are generated according to an AR(1) process with two break dates. The results indicate that the criteria BIC and YIC do not precisely estimate the number of breaks, and the ones having heavy penalty (SIC and MIC) have a tendency to underestimate the number of changes when the break size associated to the autoregressive coefficient is small and whatever the magnitude of change for the level.
- The time series are highly correlated inside the segments since the coefficient on the lagged dependent variable has a nearly unit root. This can be explained by the fact that the series are characterized by the presence of high correlation in their structure. This is checked by the Q(15) autocorrelation test of Ljung and Box (1978) since the p-values are all equal to zero indicating that the null hypothesis of independence is rejected. Note that the test also indicates that the series are all correlated inside the segments.
- A look at the graphs of the series suggests that the dominant feature is the presence of trends, which reverse themselves for some series. The break models appear to be fitting

[21] The observed value of the *WD* max statistic is 12.499 and the corresponding critical value is 11.71.

Table 7
Estimate results for the output (1869–1987)

		ε=0.10	M=5	T=119		
Tests[a]	sup F_T (1)	sup F_T (2)	sup F_T (3)	sup F_T (4)	sup F_T (5)	UD max
	9.409	7.151	10.221*	8.274**	6.992	10.221
	WD max	sup F_T (2|1)	sup F_T (3|2)	sup F_T (4|3)	sup F_T (5|4)	
	13.477	11.003	2.399	3.089	3.089	
	SIC	BIC	YIC	MIC	Seq	Rep
Number of breaks	0	0	3	0	0	0
Estimates with one Break	$\hat{T}_1$					
	1933 (1921–1938)					
	$\hat{\delta}_{1,1}$	$\hat{\delta}_{2,1}$	$\hat{\delta}_{1,2}$	$\hat{\delta}_{2,2}$		
	0.068 (0.036)	0.965 (0.021)	0.131 (0.076)	0.993 (0.016)		

[a] The sup $F_T(l+1|l)$ tests are not significant even at the 10% level (the corresponding critical values for l=1,. . . ,4 are, respectively, 12.19, 13.20, 13.79 and 14.37). The WD max test is significant at the 10% significance level since his observed value is 12.499 and the corresponding critical value is 11.71.

* Indicates that the tests are significant at the 5% level.

** Indicates that the sup F_T(4) test is significant at the 10% level (the corresponding critical value is 7.68).

a step function through these trends. Each step is large, and this is likely the reason for the quite small 95% confidence intervals associated with the estimated break dates.

- The reported results show that the selection procedures based on the tests for multiple structural breaks are more powerful than the model selection criteria in detecting changes in the level and the persistence of the series. In the same context, Jouini and Boutahar (2003) show that even for small jump size of the break date, the Bai and Perron's sequential procedure can detect breaks.
- We conclude that we can be confident that the adopted AR(1) process appears to be a good description of the time series since it provides significant results for some procedures, leads to an adequate and precise number of breaks selected by the sequential procedure and the $F_T(l+1|l)$ test and allows to well capture the breaks since their locations coincide with important international economic events.
- With the same conditions imposed on the distributions of the regressors and the errors over segments, we have determined the estimate results when we run regressions with change in mean, i.e. $\boldsymbol{x}_t$=0 and $\boldsymbol{z}_t$=1 (the results are available upon request from the authors) which means that we account for potential serial correlation via nonparametric correction. The results indicate that all the procedures do not lead to satisfactory choices of the number of structural breaks and select the maximum possible number on the overwhelming majority of cases. We would have to consider the possibility that this is explained by the fact that the series are characterized by the presence of high correlation in their structure. Hence, this highlights the practical importance of the simulation results of Perron (1997) who shows that the criteria BIC and MIC have tendency to overestimate the number of changes for an AR(1) process with high correlation degree and when we estimate a model with change in mean. The results also indicate that, in addition to the information criteria, the procedures based on the tests also select the maximum permitted number of breaks. When the dynamics of the series are taken into account in a parametric way (i.e. $\boldsymbol{x}_t$=y_{t-1} and $\boldsymbol{z}_t$=1), the empirical results are improved since the selection procedures do not lead to the overestimation of the number of structural breaks, but they globally remain inadequate since these procedures do not precisely choose the number of breaks and often have a tendency to underestimate.[22] Note that the estimates of the autoregressive coefficient approach unity explaining the problem of selecting a high number of changes when the correlation is not explicitly modelled in a parametric way.

7. Conclusion

In this paper, the question of the instability has been subjected to a meticulous examination based on some recent developments in the analysis of structural change models. Indeed, we have discussed test statistics, selection of the number of breaks and

[22] This proves the simulation results of Perron (1997) who shows that when the autoregressive coefficient is large, the criterion MIC performs better under the null hypothesis of no structural change but underestimates the number of breaks when some are present.

their locations. We have illustrated the usefulness of the procedures through a few empirical applications to U.S. time series. The obtained results are significant since the dates of the breaks coincide with important international economic events. Indeed, they may be linked to major events in the International Monetary System, the two Oil-Price Shocks and international economic problems. This paper is then justified by our aim to find economic explanations showing why in the selected dates there are shifts in the series. Another feature of substantial importance is that the presence of high correlation in the data greatly affects the estimation precision of some procedures.

Acknowledgements

We would like to thank Anne PEGUIN and Claude DENIAU for their helpful remarks and comments. We are also grateful to Eric GIRARDIN and the participants in the Workshop *Macroéconométrie* at the GREQAM, Marseille, May 2, 2003, the *XIIIth Journées du Séminaire d'Etudes et de Statistiques Appliquées á la Modélisation en Economie (SESAME)* at the GEMMA, Caen, September 8–10, 2003, the *European-Summer School, Comparative European Applied Econometrics: Macroeconometrics and Financial Econometrics* in Aix-en-Provence, September 23–26, 2003 and the *Journée d'Econométrie: Développements Récents de l'Econométrie Appliquée á la Finance* at the Paris X University, Paris, November 27, 2003 for their suggestions.

Appendix A. Definition and sources of the data

(1) The two postwar monthly U.S. interest rates (the federal funds rate and the federal discount rate) are available from January 1957 to October 2002 (yielding 550 observations) and obtained from the Federal Reserve Board of Governors.
(2) The monthly nominal exchange rates euro/U.S. dollar and yen/U.S. dollar cover, respectively, the periods 1980:1–2001:5 (257 observations) and 1971:1–2003:4 (388 observations) and are obtained from the St. Louis Reserve Federal Bank database.
(3) The annual U.S. Gross National Product (GNP) series covers the period 1869–1987 and the readers are referred to Hansen (1992b) for details on the definition and source of the series.

Appendix B. Graphics of the Series, the Residual Variance and the Wald Test Sequences as a Function of Break Date

(1) Top left: Graphics of the series.
(2) Top right: Graphics of the residual variance as a function of break date.
(3) Bottom left: Graphics of the Wald test sequence constructed using a heteroskedasticity-consistent covariance matrix for the residuals as a function of break date.

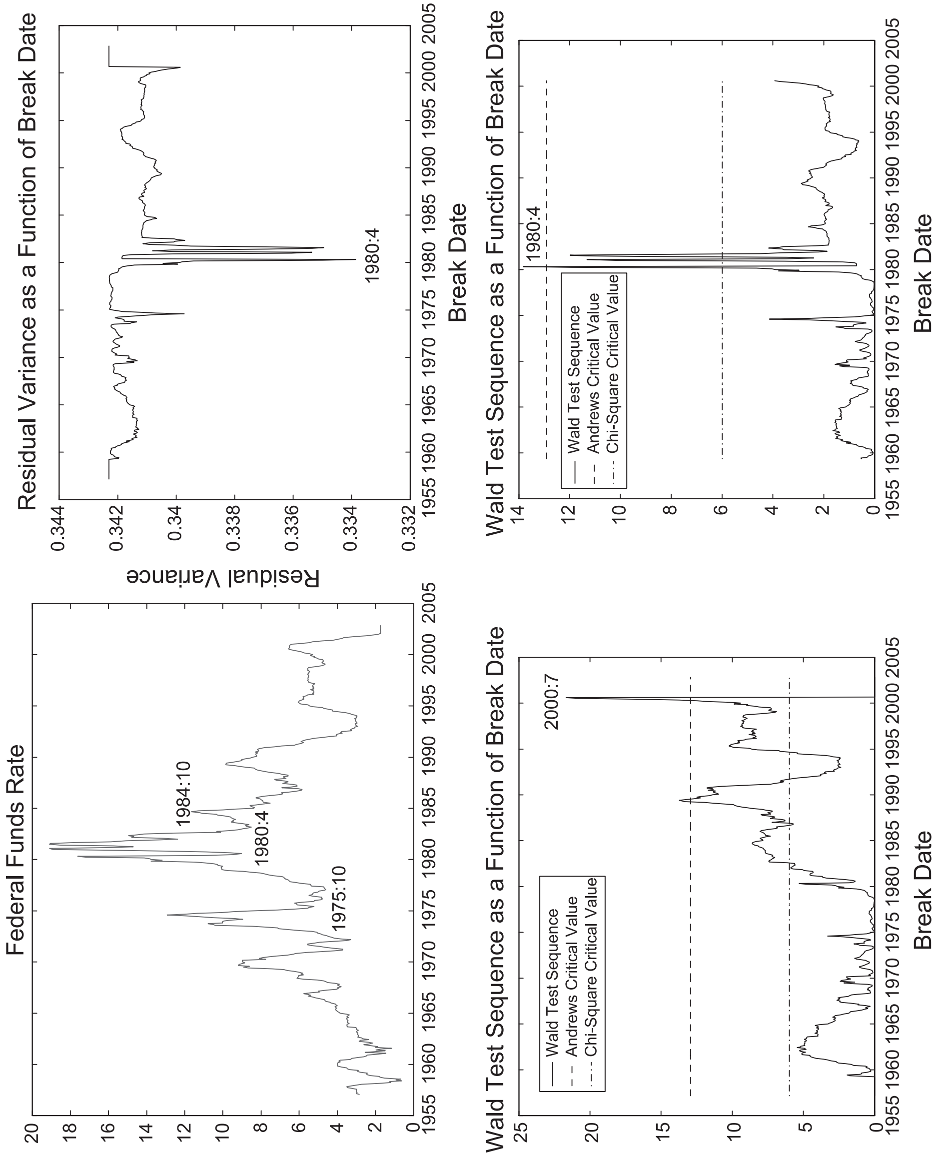

Fig. 1. Graphics for the U.S. federal funds rate.

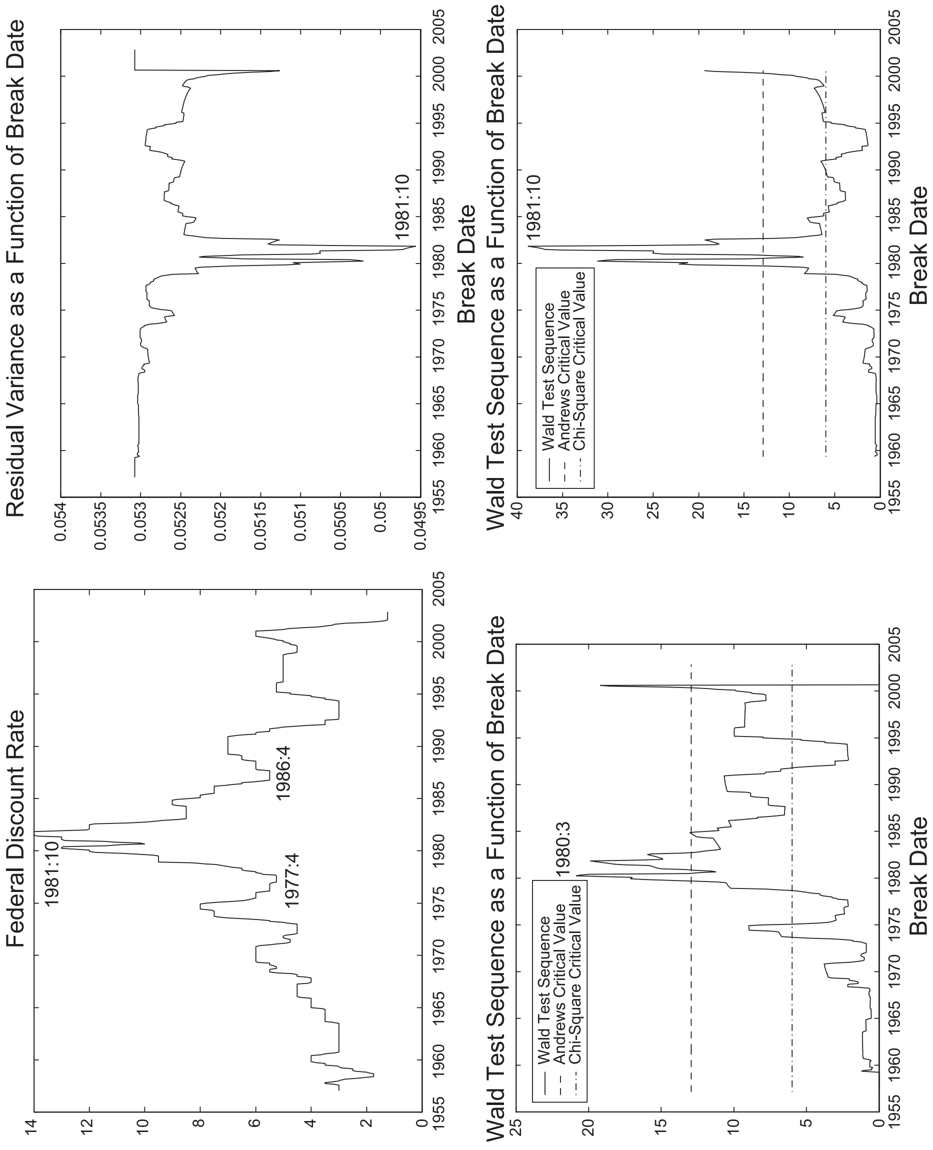

Fig. 2. Graphics for the U.S. federal discount rate.

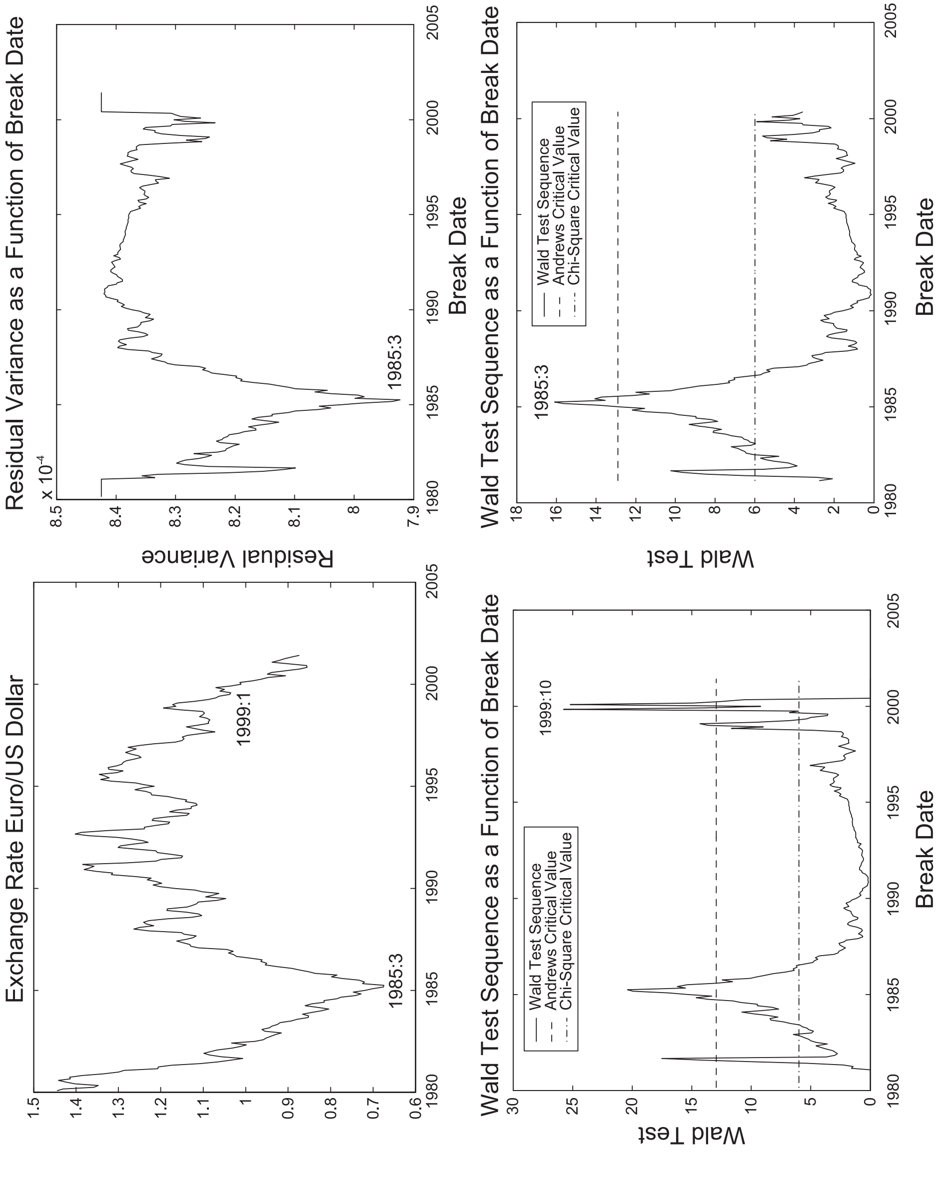

Fig. 3. Graphics for the exchange rate euro/U.S. dollar.

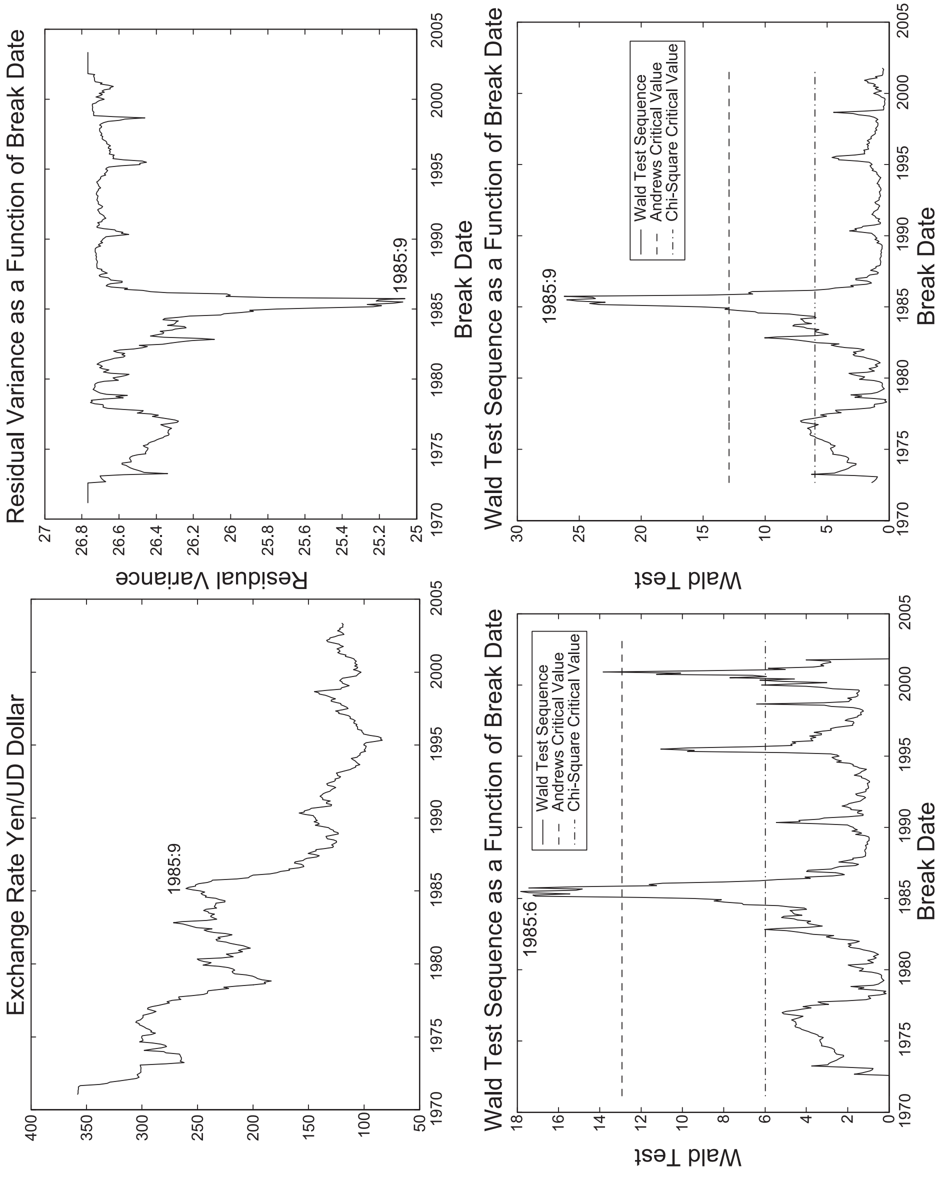

Fig. 4. Graphics for the exchange rate yen/U.S. dollar.

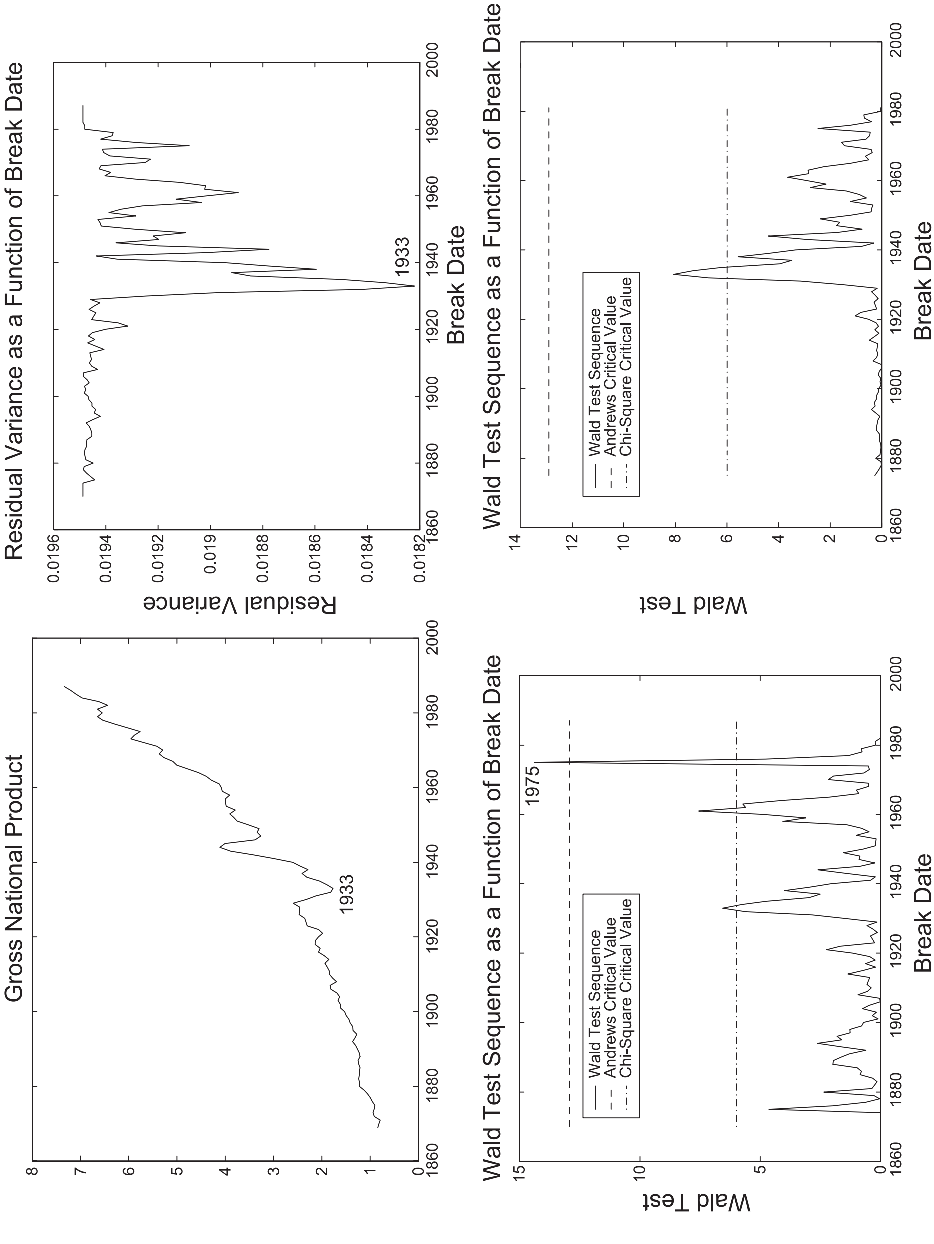

Fig. 5. Graphics for the U.S. gross national product.

(4) Bottom right: Graphics of the Wald test sequence constructed using the homoskedastic form of the covariance matrix for the residuals as a function of break date.

References

Alogoskoufis, G.S., Smith, R., 1991. The Phillips curve, the persistence of inflation, and the Lucas critique: evidence from exchange rate regimes. American Economic Review 81, 1254–1275.

Andrews, D.W.K., 1993. Tests for parameter instability and structural change with unknown change point. Econometrica 61, 821–856.

Andrews, D.W.K., Ploberger, W., 1994. Optimal tests when a nuisance parameter is present only under the alternative. Econometrica 62, 1383–1414.

Bai, J., 1994. Least squares estimation of a shift in linear processes. Journal of Time Series Analysis 15, 453–472.

Bai, J., 1997a. Estimation of a change point in multiple regression models. Review of Economics and Statistics 79, 551–563.

Bai, J., 1997b. Estimating multiple breaks one at a time. Econometric Theory 13, 315–352.

Bai, J., 1999. Likelihood ratio tests for multiple structural changes. Journal of Econometrics 91, 299–323.

Bai, J., Perron, P., 1998. Estimating and testing linear models with multiple structural changes. Econometrica 66, 47–78.

Bai, J., Perron, P., 2003a. Computation and analysis of multiple structural change models. Journal of Applied Econometrics 18, 1–22.

Bai, J., Perron, P., 2003b. Critical values for multiple structural change tests. Econometrics Journal 1, 1–7.

Bai, J., Perron, P., 2004. Multiple structural change models: a simulation analysis. In: Corbea, D., Durlauf, S., Hansen, B.E. (Eds.), Econometric Essays. Cambridge University Press. In press.

Bai, J., Lumsdaine, R.L., Stock, J.H., 1998. Testing for and dating common breaks in multivariate time series. Review of Economic Studies 64, 395–432.

Ben Aïssa, M.S., Jouini, J., 2003. Structural breaks in the US inflation process. Applied Economics Letters 10, 633–636.

Ben Aïssa, M.S., Boutahar, M., Jouini, J., 2004. Bai and Perron's and spectral density methods for structural change detection in the US inflation process. Applied Economics Letters 11, 109–115.

Boutahar, M., Jouini, J., 2003a. Spuriousness of information criteria in selecting the number of changes: theoretical study and simulation evidence, Working Paper n° 03A17, GREQAM, University of Méditerranée, Marseille.

Boutahar, M., Jouini, J., 2003b. Is the US inflation process nonstationary and mean-reverting? A fractional integration investigation, Paper presented at the International Workshop on Models with Breaks in Economics and Finance: Recent Developments, November, 2003, Paris.

Campbell, J.Y., Shiller, R.J., 1987. Cointegration and tests of present value models. Journal of Political Economy 95, 1062–1088.

Campbell, J.Y., Shiller, R.J., 1988a. Stock prices, earning and expected dividends. The Journal of Finance 43, 661–676.

Campbell, J.Y., Shiller, R.J., 1988b. The dividend-price ratio and expectations of future dividends and discount factors. Review of Financial Studies 1, 195–228.

Chow, G.C., 1960. Tests of equality between sets of coefficients in two linear regressions. Econometrica 28, 591–605.

Chong, T.T.L., 1994. Consistency of change point estimators when the number of change points in structural change models is underspecified, Unpublished Manuscript, Department of Economics, University of Rochester.

Chong, T.T.L., 1995. Partial parameter consistency in a misspecified structural change model. Economics Letters 49, 351–357.

Chong, T.T.L., 2001. Structural change in AR(1) models. Econometric Theory 17, 87–155.

Garcia, R., Perron, P., 1996. An analysis of the real interest rate under regime shifts. Review of Economics and Statistics 78, 111–125.

Hansen, B.E., 1992a. Tests for parameter instability in regressions with I(1) processes. Journal of Business and Economic Statistics 10, 321–335.
Hansen, B.E., 1992b. Testing for parameter instability in linear models. Journal of Policy Modeling 14, 517–533.
Hansen, B.E., 1997. Approximate asymptotic p-values for structural change tests. Journal of Business and Economic Statistics 15, 60–67.
Hansen, B.E., 2000. Testing for structural change in conditional models. Journal of Econometrics 97, 93–115.
Hansen, B.E., 2001. The new econometrics of structural change: dating changes in US labor productivity. Journal of Economic Perspectives 15, 117–128.
Hawkins, D.M., 1977. Testing a sequence of observations for a shift in location. Journal of the American Statistical Association 72, 180–186.
James, B., James, K., Siegmund, D., 1987. Tests for a change point. Biometrika 74, 71–83.
Jouini, J., Boutahar, M., 2003. Structural breaks in the US inflation process: a further investigation. Applied Economics Letters 10, 985–988.
Liu, J., Wu, S., Zidek, J.V., 1997. On segmented multivariate regressions. Statistica Sinica 7, 497–525.
Ljung, G.M., Box, G.E.P., 1978. On a measure of lack of fit in time series models. Biometrika 65, 553–564.
Mankiw, N.G., Miron, J.A., 1986. The changing behavior of the term structure of interest rates. Quarterly Journal of Economics 101, 221–228.
Mankiw, N.G., Miron, J.A., Weil, D.N., 1987. The adjustment of expectations to a change in regime: a study of the founding of the Federal Reserve. American Economic Review 77, 358–374.
Nabeya, S., Tanaka, K., 1988. Asymptotic theory of a test for the constancy of regression coefficients against the random walk alternative. Annals of Statistics 16, 218–235.
Nunes, L.C., Newbold, P., Kuan, C.M., 1996. Spurious number of breaks. Economics Letters 50, 175–178.
Nyblom, J., 1989. Testing for the constancy of parameters over time. Journal of the American Statistical Association 84, 223–230.
Perron, P., 1989. The great crash, the oil-price shock, and the unit-root hypothesis. Econometrica 57, 1361–1401.
Perron, P., 1997. L'estimation de modéles avec changements structurels multiples. Actualité Economique 73, 457–505.
Quandt, R.E., 1958. The estimation of parameters of a linear regression system obeying two separate regimes. Journal of the American Statistical Association 55, 873–880.
Quandt, R.E., 1960. Tests of the hypothesis that a linear regression obeys two separate regimes. Journal of the American Statistical Association 55, 324–330.
Schwarz, G., 1978. Estimating the dimension of a model. Annals of Statistics 6, 461–464.
Sen, A., Srivastava, M.S., 1975a. On tests for detecting change in mean. Annals of Statistics 3, 96–103.
Sen, A., Srivastava, M.S., 1975b. Some one-sided tests for change in level. Technometrics 17, 61–64.
Srivastava, M.S., Worsley, K.J., 1986. Likelihood ratio tests for a change in the multivariate normal mean. Journal of the American Statistical Association 81, 199–204.
Worsley, K.J., 1979. On the likelihood ratio test for a shift in locations of normal populations. Journal of the American Statistical Association 72, 180–186.
Worsley, K.J., 1986. Confidence regions and tests for a change-point in a sequence of exponential family random variables. Biometrika 73, 91–104.
Yao, Y.C., 1988. Estimating the number of change-points via Schwarz' criterion. Statistics & Probability Letters 6, 181–189.
Yao, Y.C., Au, S.T., 1989. Least squares estimation of a step function. Sankhyā. Series A 51, 370–381.
Yin, Y.Q., 1988. Detection of the number, locations and magnitudes of jumps. Communications in Statistics. Stochastic Models 4, 445–455.
Zivot, E., Andrews, D.W.K., 1992. Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. Journal of Business and Economic Statistics 10, 251–270.